Information Sciences Letters

Volume 10
Issue 3 *Sep. 2021*

Article 19

2021

In-silico Analysis of Protein Receptors Contributing to SARS-COV-2 High Infectivity

Ismail Atia
Zoology Department, Faculty of Science, Al–Azhar University, Assiut 71524, Egypt\\ Center of Excellence in Cancer Research, Tanta University Teaching Hospital, Tanta, Egypt, dr.ismail_atia@yahoo.com

Mohamed L Salem
Department of Zoology, Faculty of Science, Tanta University, Tanta, Egypt\\ Center of Excellence in Cancer Research, Tanta University Teaching Hospital, Tanta, Egypt, dr.ismail_atia@yahoo.com

Aya Elkholy
Faculty of Applied Medical Science, Pharos University, Alex, Egypt, dr.ismail_atia@yahoo.com

Wael Elmashad
Physiology Department, Faculty of Medicine, Tanta University, Egypt, dr.ismail_atia@yahoo.com

Gomaa A. M. Ali
Chemistry Department, Faculty of Science, Al–Azhar University, Assiut 71524, Egypt, dr.ismail_atia@yahoo.com

Inf. Sci. Lett. **10**, No. 3, 561-570 (2021)

Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/100320

In-silico Analysis of Protein Receptors Contributing to SARS-COV-2 High Infectivity

Ismail Atia [1,3,*], *Mohamed L Salem* [2,3], *Aya Elkholy* [4], *Wael Elmashad* [5] *and Gomaa A.M. Ali* [6]

[1]Zoology Department, Faculty of Science, Al–Azhar University, Assiut 71524, Egypt
[2]Department of Zoology, Faculty of Science, Tanta University, Tanta, Egypt
[3]Center of Excellence in Cancer Research, Tanta University Teaching Hospital, Tanta, Egypt
[4]Faculty of Applied Medical Science, Pharos University, Alex, Egypt
[5]Physiology Department, Faculty of Medicine, Tanta University, Egypt
[6]Chemistry Department, Faculty of Science, Al–Azhar University, Assiut 71524, Egypt

Received: 19 Jun. 2021, Revised: 2 Jul. 2021; Accepted: 1 Aug. 2021
Published online: 1 Sep. 2021.

Abstract: SARS-CoV-2 attacked more than 120 million people and causing the death of more than two million worldwide. Because of the crucial role of ACE2 protein as an entry for SRS-COV2, we investigated the protein's sequence in seventy-three living species. Data analysis of protein sequences, ACE2 mRNA, expression analysis, and protein interaction for humans and other living species were obtained from databases. The phylogenetic tree was constructed using MEGA6. We found 95% or more similarity between the conserved protein domains between Homo sapiens and Felis catus, Pan troglodytes, Pan paniscus, and Equus caballus. These species could be expressed the protein in their cell surface with the same properties as Homo sapiens. This leads to the idea of being an actual transmitter of the virus SARS-COV2, and maybe a possible reason for the spread of the virus when work or play with it, eating, cooking it, or transfer from one place to another. Expression analyses provide more explanations about organs in the body that expressed more genes like lung, heart, small intestine, and colon, which are affected more than other organs or tissues during infection or are supposed to be an infection transmitter when dealing with it in the animal after sacrifices or die. We concluded that the possibility of high SARS-CoV-2 infectivity via both zoonosis and reverse zoonosis is interesting and needs more research to develop a new strategy for dealing with this virus.
Keyword: COVID-19, SARS-CoV-2, Reverse zoonosis, ACE2 mRNA, Phylogenetic tree.

1 Introduction

At the end of 2019, a new strain of coronaviruses (COVID-19) attacks Wuhan's city, outbreaks to other places of China, and then to other parts of the world [1]. Coronavirus is a serious pandemic due to the high infectivity rates and the wide distribution in different countries [2,3]. The high infectivity of the COVID-19 is due to the virus's ability to attach to the cell's surface and ACE2. ACE2 is an enzyme connected to lung type II alveolar cell membranes, enterocytes of the small intestine, and smooth muscle cells in most organs [4,5]. The expression of ACE2 in cortical neurons, Neurotransmitters, and glia makes them very vulnerable to SARS-CoV-2 attack, which was the possible basis of anosmia or losing his sense of taste incidences of neurological deficits seen in COVID-19 [6]. ACE2 mRNA expression is observed in the brainstem, striatum, cerebral cortex, and hypothalamus. ACE2 serves as the entry point into cells for some coronaviruses HCoV-NL63, SARS-CoV and SARS-CoV-2. More precisely, The binding of the spike S1 protein of SARS-CoV and SARS-CoV-2 to the enzymatic domain of ACE2 on the cell membrane causes phagocytosis and the virus's immobilization of the enzyme into endosomes [7,8]. The structural basis of the binding between the viral S protein and ACE2 has been significantly investigated. The critical portion of the spike glycoprotein (known as Receptor-Binding Domain or RBD) has revealed 6 critical residues (Leu455, Phe486, Gln493, Ser494, Asn501, Tyr505) at the interface with ACE26 [9].

The receptor ACE2 in the SARS-CoV-2 causes the infection of the type II pneumocyte population of the lung cells as a possible mechanism for viral infections has been identified [10]. The expression of ACE2 with increased vulnerability to viral entry was linked [7,11]. Moreover, the ACE2 expression information at the single-cell level to rank the cells based on their vulnerability to infection with SARS-CoV-2 has been utilized. Consistent with the scRNA sequence data, the dual

*Corresponding author e-mail: dr.ismail_atia@yahoo.com

inhibition of host cell cysteine and serine proteases hindered viral entry into the cell [9]. In addition, the neuropilin-1 (NRP1) may act as a host cofactor and enable viral entry [12]. Furthermore, a unique characteristic of the human SARS-CoV-2 coronavirus that may have improved its contagiousness is the insertion of 12-nucleotides resulting in 4 different amino acids (Pro-Arg-Arg-Ala) in position 681-68410; these 4 extra residues create a polybasic cleavage site unique to this human virus (not found in any other species nor the 2002-2003 SARS-CoV virus) which may be involved in the cleavage of the spike protein, facilitating entry in target cells. Multispecies sequence alignment at the nucleotide and amino acid level is a valuable technique to reconstruct the phylogenesis of viruses or receptor hosts, allowing us to trace the probable origin of the SARS-CoV-2 to the source viruses and defined the probable reservoir host or the transference vectors for the SRS-COV-2.

To investigate the importance of these protein receptors in the defined living species that contributed to the virus transmission process from animal to a human called (zoonosis), human-to-animal (reverse zoonosis), or transferred from animal-to-animal. This study identified the ACE2 protein receptor in seventy-three living species and constructed phylogenetic trees using the protein sequences, conserved protein domains, and mRNA sequences elucidated to examine these genes in humans compared to other living species. This analysis may open an avenue for more research about the infectivity of SRS-COV2 in the future.

2 Material and Methods

2.1 Data Resources

The ACE2 protein sequences for human and other living species used for sequence alignment analysis and phylogenetic tree construction were obtained from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein). Protein conserved domains were obtained from the NCBI CCD tool from the NCBI database (http://www.ncbi.nlm.nih.gov/CCD). mRNA sequences were obtained from the mRNA tool from the NCBI database (http://www.ncbi.nlm.nih.gov/mRNA). Datasets for expression analysis were obtained from BioGPS database (http://biogps.org/#goto). Protein interaction network was obtained from Gene Card human gene database (http://www.genecards.org/). The phylogenetic tree was constructed using the Molecular Evolutionary Genetics Analysis program6 (MEGA6) [13].

2.2 Identification of Protein Sequences, Alignment, and Construction of the Phylogenetic Tree

ACE2 protein sequences for 73 different living species were obtained from NCBI database and used as a template for analysis. The alignments were analyzed using the Clustal W sequence alignment method. All algorithms are used without additional software packages and on all major platforms. An R interface complements multiple sequence algorithms to the powerful LaTeX package text shade allows for a highly customizable plot of multiple sequence alignments [14-16], and phylogenetic trees were assembled using MEGA6 by the neighbor-joining approach for branching points in a tree using an approach that is based on the RelTime method (RelTime is described elsewhere [17]) which does not require assumptions for lineage rate variations. The implementation in MEGA is very fast and expands on the RelTime method so that multiple calibration constraints can be provided, in which case MEGA will produce absolute divergence times and relative divergence times while respecting the provided constraints. Additionally, the implementation in MEGA can compute divergence times without calibration constraints, in which case, only relative times will be produced [18-20].

2.3 mRNA Sequences Identification and Protein Domains Analysis

The various organisms' functional protein domains were obtained from the NCBI database and analyzed using the protein family database (Pfam) sequence tool [21], shown in Figure Version = cdd. v.3.19, Info Source: Precalculated Data, Preset Options: CD-SEARCH/cdd database Filter of low complexity: no Composition Based Adjustment is allowed, and the E-value threshold is set to 0.01. The mRNA sequences of each gene that contributed to the translation of the ACE2 in the needed species were obtained from the NCBI database's mRNA; then, the sequences were aligned to construct the phylogenetic tree of mRNA sequences [22,23].

3 Results

3.1 Identification of Protein Sequences in the Species and Establishing of the Phylogenetic Tree

To characterize ACE2 protein, we investigated the protein sequence in seventy-three living species, including Homo sapiens, Mus musculus Felis catus Rattus norvegicus, Bos Taurus, Equus caballus, Pan troglodytes, Canis lupus familiaris, Cavia porcellus, Ovis aries, Macaca mulatta.
Mesocricetus auratus, Capra hircus, Sus scrofa domesticus, Procyon lotor, Sus scrofa, Oryctolagus cuniculus, Columba livia, Gallus gallus, Paguma larvata, Pongo abelii, Pan troglodytes, Mustela putorius furo, Ornithorhynchus anatinus, Pongo abelii, Caenorhabditis elegans, Nomascus leucogenys, Anolis carolinensis, Danio rerio,Monodelphis domestica, Cercocebus atys, Macaca nemestrina, Mandrillus, leucophaeus, Ictidomys tridecemlineatus, Chlorocebus sabaeus, Macaca fascicularis, Myotis lucifugus, Pan paniscus, Papio Anubis, Loxodonta Africana, Rhinopithecus roxellana, Heterocephalus glaber, Propithecus coquereli, Ursus maritimus, Macaca mulatta, Ficedula albicollis, tolemur garnettii, Saimiri boliviensis boliviensis, Meleagris

gallopavo, Mesocricetus auratus, Cebus capucinus imitator, Ailuropoda melanoleuca, Dipodomys ordii, Vombatus ursinus, Bos indicus Taurus, Tarsius syrichta, Pelodiscus sinensis, Rousettus leschenaultia, Nyctereutes procyonoides, Rhinolophus ferrumequinum, Aotus nancymaae, Vulpes vulpes, Ursus americanus, Physeter macrocephalus, Physeter macrocephalus, Rhinopithecus bieti, Anas platyrhynchos platyrhynchos, Rhinolophus sinicus, Rousettus leschenaultii, Taeniopygia guttata, Alligator sinensis, Electrophorus electricus, Colobus angolensis palliates and Latimeria *chalumnae* representing animals from Nematodes, fishes, reptiles, aves and mammals which explained in their common names and accession numbers codes obtained from NCBI database (Table 1).

The MEGA6 program further analyzed these proteins' function and conservation and explained the evaluation between species used to analyze these proteins' sequences. The constructed phylogenetic tree for twenty of the organisms and explains that the protein structure of *Homo sapiens* has a relationship with some of the other protein sequences starts with the very closed relation like *Pan troglodytes,* which share with *Homo sapiens* the same sequence, and *Macaca mulatta,* which is very close to *Homo sapiens* in the same tree part. Differences between these three organisms seem to disappear. On the other hand, in a nearby part of the tree Felis catus, Canis lupus familiaris and Procyon lotor have evolutionary distance close to the tree's first part. *Bos Taurus, Ovis aries, Capra hircus, Sus scrofa domesticus and Equus caballus* have a closed structural distance from the *Homo sapiens* in the tree (Fig. 1).

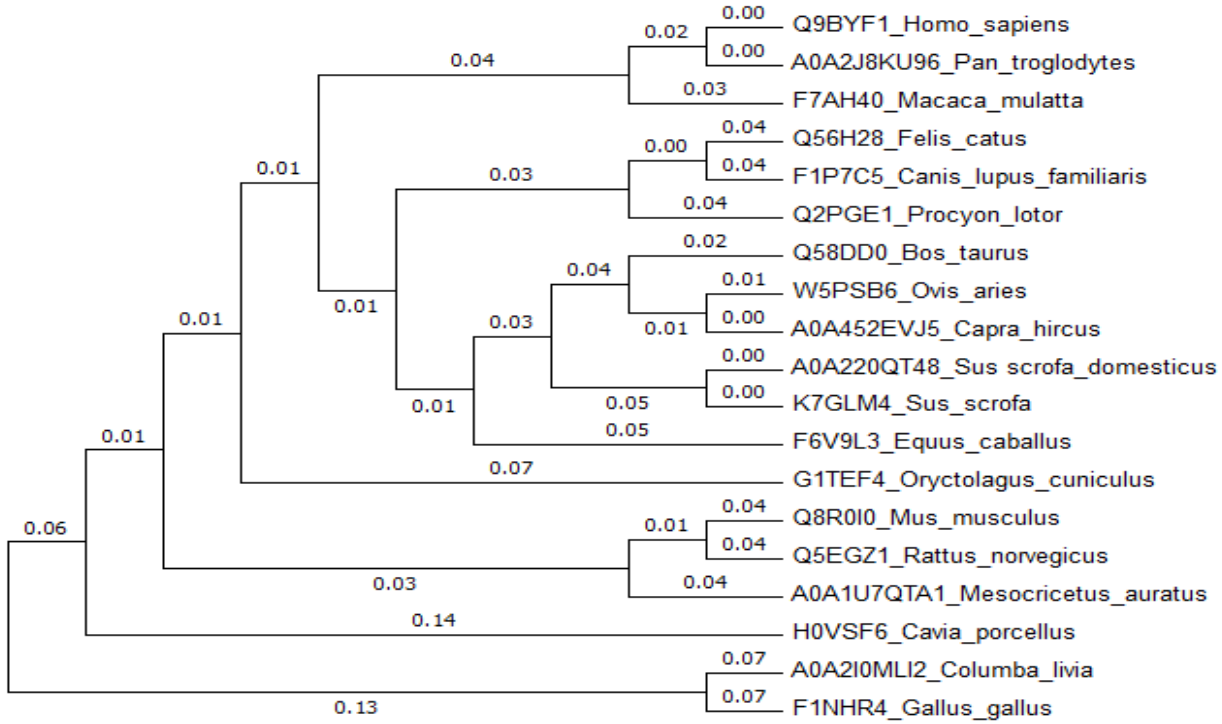
Table 1. Data representing animals from Nematodes, reptiles, aves, and mammals are explained in their scientific names, common names, and accession numbers codes obtained from the NCBI database.

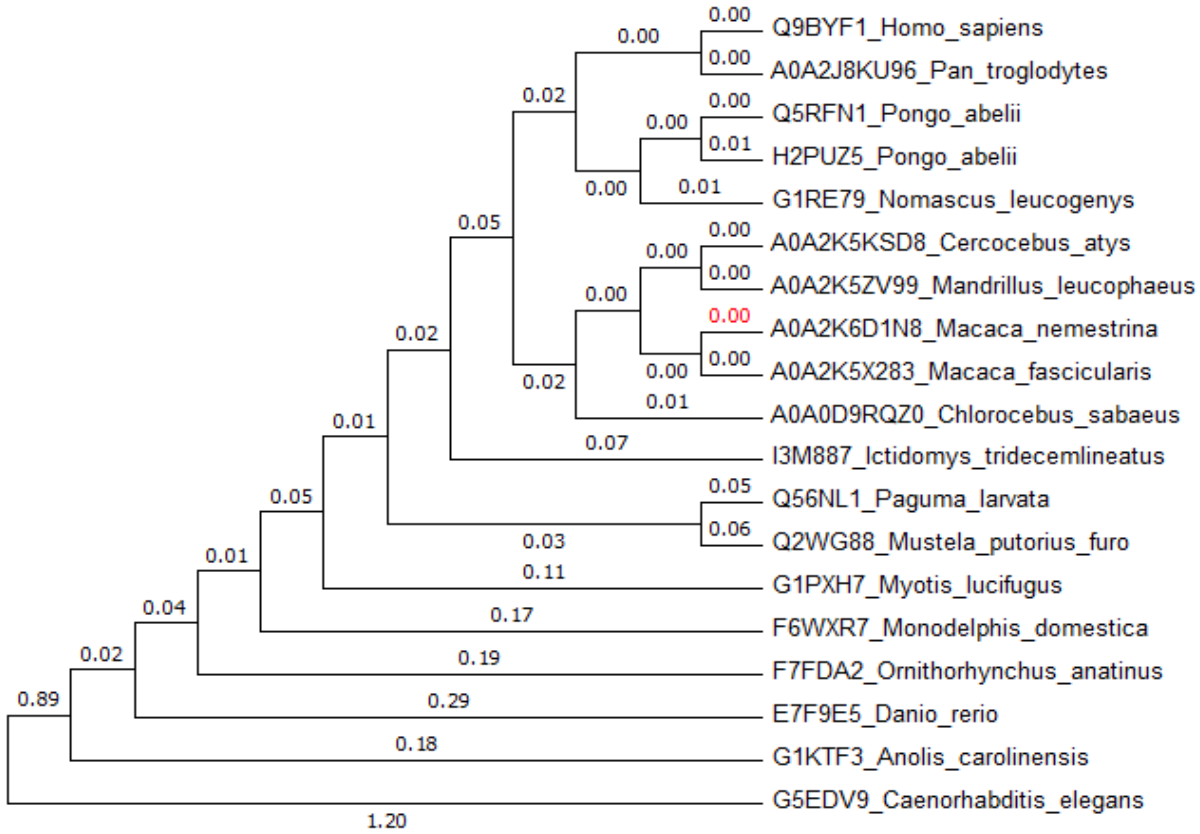
No.	**Accession number**	**Scientific name**	**Common name**
1	Q9BYF1	*Homo sapiens*	Human
2	Q8R0I0	*Mus musculus*	Mouse
3	Q5EGZ1	*Rattus norvegicus*	Brown rat
4	Q56H28	*Felis catus*	domestic cat
5	Q58DD0	*Bos taurus*	Cow
6	F6V9L3	*Equus caballus*	Horse
7	A0A2J8KU96	*Pan troglodytes*	Chimpanzee
8	F1P7C5	*Canis lupus familiaris*	Dog
9	H0VSF6	*Cavia porcellus*	Guinea pig
10	W5PSB6	*Ovis aries*	Sheep
11	F7AH40	*Macaca mulatta*	Rhesus macaque
12	A0A1U7QTA1	*Mesocricetus auratus*	Golden hamster
13	A0A452EVJ5	*Capra hircus*	*Goat*
14	A0A220QT48	*Sus scrofa domesticus*	*domestic pig*
15	Q2PGE1	*Procyon lotor*	*Raccoon*
16	K7GLM4	*Sus scrofa*	*Pig*
17	G1TEF4	*Oryctolagus cuniculus*	*Rabbit*
18	A0A2I0MLI2	*Columba livia*	Rock dove
19	F1NHR4	*Gallus gallus*	The red junglefowl
20	Q56NL1	*Paguma larvata*	Masked palm civet
21	Q5RFN1	*Pongo abelii*	The Sumatran orangutan
22	A0A2J8KU96	*Pan troglodytes*	Chimpanzee
23	Q2WG88	*Mustela putorius furo*	European domestic ferret
24	F7FDA2	*Ornithorhynchus anatinus*	The platypus
25	H2PUZ5	*Pongo abelii*	Sumatran orangutan
26	G5EDV9	*Caenorhabditis elegans*	Cae-elegans
27	G1RE79	*Nomascus leucogenys*	Northern white-cheeked gibbon
28	G1KTF3	*Anolis carolinensis*	Green American chameleon
29	E7F9E5	*Danio rerio*	Zebrafish
30	F6WXR7	*Monodelphis domestica*	Gray short-tailed opossum
31	A0A2K5KSD8	*Cercocebus atys*	Sooty mangabey
32	A0A2K6D1N8	*Macaca nemestrina*	Pig-tailed macaque
33	A0A2K5ZV99	*Mandrillus leucophaeus*	Drill
34	I3M887	*Ictidomys tridecemlineatus*	Thirteen-lined ground squirrel
35	A0A0D9RQZ0	*Chlorocebus sabaeus*	Green monkey
36	A0A2K5X283	*Macaca fascicularis*	Crab-eating macaque
37	G1PXH7	*Myotis lucifugus*	Little brown bat
38	A0A2R9BKD8	*Pan paniscus*	Pygmy chimpanzee
39	A0A096N4X9	*Papio anubis*	Olive baboon
40	G3T6Q2	*Loxodonta africana*	African elephant
41	A0A2K6NFG7	*Rhinopithecus roxellana*	Golden snub-nosed monkey

42	A0A0N8EUX7	*Heterocephalus glaber*	Naked mole rat
43	A0A2K6GHW5	*Propithecus coquereli*	Coquerel's sifaka
44	A0A452TT30	*Ursus maritimus*	Polar bear
45	F7AH40	*Macaca mulatta*	Rhesus macaque
46	U3JP73	*Ficedula albicollis*	Collared flycatcher
47	H0WMI5	*tolemur garnettii*	Small-eared galago
48	A0A2K6SBD4	*Saimiri boliviensis boliviensis*	Bolivian squirrel monkey
49	G1NPB8	*Meleagris gallopavo*	Wild turkey
50	A0A1U7QTA1	*Mesocricetus auratus*	Golden hamster
51	A0A2K5PYM0	*Cebus capucinus imitator*	Panamanian white-faced capuchin
52	G1MC42	*Ailuropoda melanoleuca*	Giant panda
53	A0A1S3GHT7	*Dipodomys ordii*	Ord's kangaroo rat
54	A0A4X2M679	*Vombatus ursinus*	Common wombat
55	A0A4W2H6E0	*Bos indicus taurus*	Hybrid cattle
56	A0A1U7TY97	*Tarsius syrichta*	Philippine tarsier
57	K7FJ41	*Pelodiscus sinensis*	Chinese softshell turtle
58	A4PIG8	*Rousettus leschenaultii*	Leschenault's rousette
59	B4XEP4	*Nyctereutes procyonoides*	Raccoon dog
60	E2DHI2	*Rhinolophus ferrumequinum*	Greater horseshoe bat
61	A0A2K5DQI6	*Aotus nancymaae*	Ma's night monkey
62	A0A3Q7RAT9	*Vulpes vulpes*	Red fox
63	A0A452R1Z9	*Ursus americanus*	American black bear
64	A0A2Y9S5T9	*Physeter macrocephalus*	Sperm whale
65	A0A2K6LKA0	*Rhinopithecus bieti*	Black snub-nosed monkey
66	U3J4G2	*Anas platyrhynchos platyrhynchos*	Northern mallard
67	U5WHY8	*Rhinolophus sinicus*	Chinese rufous horseshoe bat
68	A4PIG8	*Rousettus leschenaultii*	Leschenault's rousette
69	H0ZCK6	*Taeniopygia guttata*	Zebra finch
70	A0A3Q0H852	*Alligator sinensis*	Chinese alligator
71	A0A4W4EE33	*Electrophorus electricus*	Electric eel
72	A0A2K5JE65	*Colobus angolensis palliatus*	Peters' Angolan colobus
73	H3B2W0	*Latimeria chalumnae*	Coelacanth

The second part of the tree showed that *Pan Troglodytes* is the nearest species to the *Homo sapiens,* which seems to be no structural differences between them *Pongo abelii* and *Nomascus leucogenys* are the closest part of the tree-like *Homo sapiens*. While the third part of the tree represents *Cercocebus atys, Macaca fascicularis, Macaca nemestrina Mandrillus, leucophaeus,* and *Chlorocebus sabaeus* which is also closed to the *Homo sapiens* the other species in the tree have a large structural difference from *Homo sapiens* (Fig. 2). The third part of the phylogenetic tree showed that *Pan Paniscus* is the nearest species to *Homo sapiens, Papio Anubis, Macaca mulatta,* and *Rhinopithecus roxellana* are closed in the structural to *Homo sapiens*. In contrast, the other species have a large evolutionary distance from *Homo sapiens* (Fig. 3). The fourth part of the phylogenetic protein tree showed that *Rhinopithecus bieti* is the closest species to Homo sapiens, not very close but the closest species between the tree's other species, while all the other species have a large evolutionary distance to the *Homo sapiens* (Fig. 4).

According to the phylogenetic tree results, there are some organisms closed to Homo sapiens in the structure, which are Felis catus, Pan troglodytes, Pan paniscus Equus caballus, Bos Taurus, Canis lupus familiaris, Ovis aries, Macaca mulatta, Capra hircus, Sus scrofa domesticus, Procyon lotor, Sus scrofa, Papio Anubis, Macaca mulatta, Rhinopithecus roxellana, Saimiri boliviensis boliviensis, Cebus capucinus imitator, Rhinopithecus bieti, Colobus angolensis palliates, Piliocolobus tephrosceles, Pongo abelii, Nomascus leucogenys, Cercocebus atys, Macaca nemestrina, Mandrillus leucophaeus, Chlorocebus sabaeus and Macaca fascicularis. To get more knowledge about these species, the conserved domains were analyzed.





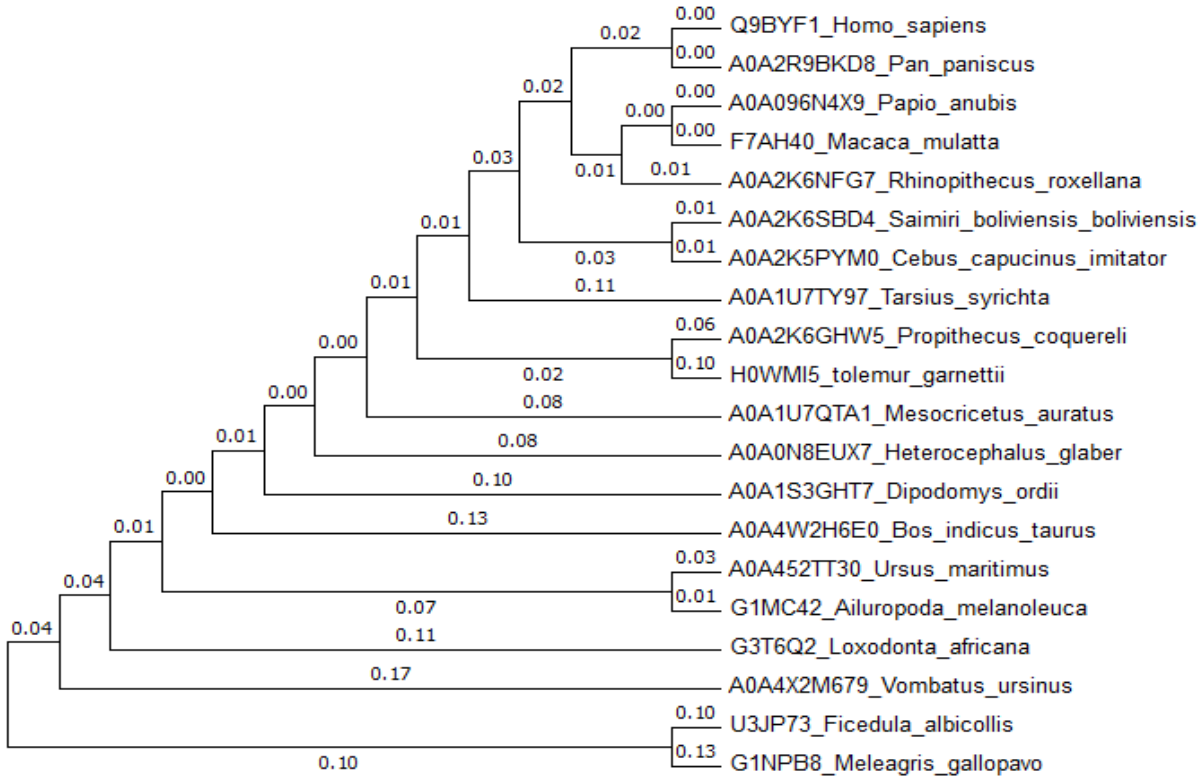


Fig. 1: Protein phylogenetic trees of transcription factors of ACE2 protein of the taxa studied compared to *Homo sapiens*. The Neighbor-Joining strategy was used to conclude the evolutionary roots. The ideal tree is seen, with a branch length number of 1.21244498 (next to the trees). The evolutionary distances were calculated using the Poisson correction method and are measured in the number of amino acid substitutions per position. The study included 19 amino acid sequences. All places with gaps and incomplete data were removed. The final dataset contains 616 locations. MEGA6 was used to perform evolutionary analyses.

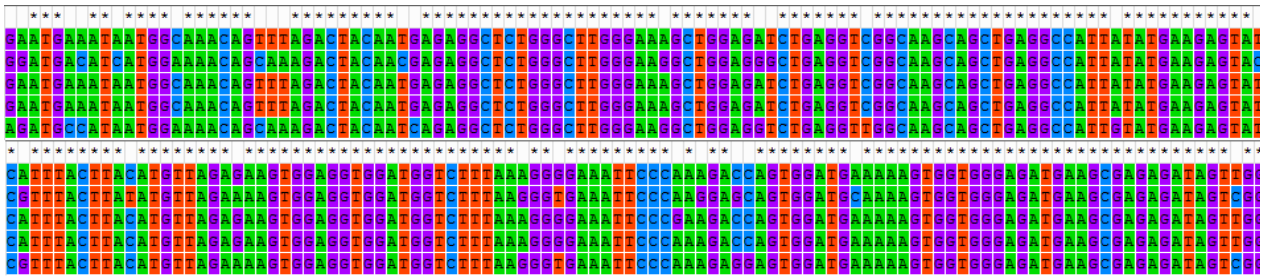


Fig. 2: Multiple alignments of mRNA sequence for the five species: Felis catus, Pan troglodytes, Pan paniscus, Equus caballus characterize this protein for the ACE2 protein of *Homo sapiens.* Each dot represents a nucleotide indicated by definite color.

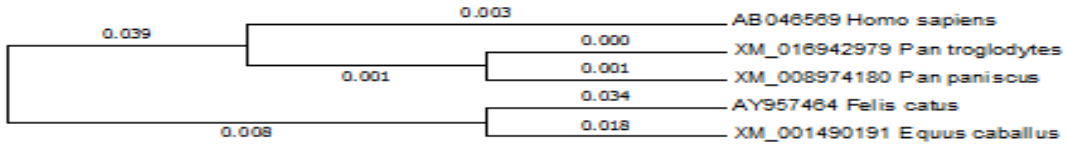


Fig. 3: Phylogenetic tree of mRNA sequences for the five species Felis catus, Pan troglodytes, Pan paniscus, Equus caballus.

3.2 Protein Domains Analysis of the Selected Species

Each protein domain is a conserved portion of a protein sequence structure that can evolve function, and it's often functional units. Analysis of ACE2 functional domains using the CCD tool for the selected species showed that *Homo sapiens* have five conserved domains: the collectrin domain (Renal amino acid transporter). Collectrin is a single-pass protein encoded that is structurally similar to the C-terminus of human angiotensin-converting enzyme 2. Peptidase_M2 domain members of this family are dipeptidyl carboxypeptidases (cleave carboxyl dipeptides), and specifically, it transforms angiotensin I to angiotensin II. (M2_ACE domain) Peptidase family M2 is a zinc-dependent

membrane-bound dipeptidase that catalyzes the transfer of the decapeptide angiotensin I to the potent vasopressor octapeptide angiotensin II by eliminating two C-terminal amino acids. (M3_like domain) The M2 angiotensin-converting enzyme (ACE, EC 3.4.15.1) is a zinc-dependent membrane-bound dipeptidase that leads to the activation of the decapeptide angiotensin I to the active vasopressor octapeptide angiotensin II and (PepF domain) Peptidase family M3 oligopeptidase F (oligendopeptidase) is majorly bacterial and includes oligoendopeptidase F from Geobacillus stearothermophilus.

Furthermore, the most similar species were *Felis catus, Pan troglodytes, Pan paniscus, and Equus caballus* share the exact five conserved domains represented in, But Felis catus has the five domains one more domain (GluZincin Peptidase family). According to Bos Taurus, Canis lupus familiaris, Ovis aries, Macaca mulatta, Capra hircus, Sus scrofa domesticus, Procyon lotor, Sus scrofa, Papio Anubis, Macaca mulatta, Rhinopithecus roxellana, Saimiri boliviensis boliviensis, Cebus capucinus imitator, Rhinopithecus bieti, Colobus angolensis palliates, Piliocolobus tephrosceles, Pongo abelii, Nomascus leucogenys, Cercocebus atys, Macaca nemestrina, Mandrillus leucophaeus, Chlorocebus sabaeus, and Macaca fascicularis we found that these species share four of the conserved domains (Collectrin, Peptidase_M2, M2_ACE, and M3_like). The analysis of these conserved domains was taken to ensure phylogenetic trees' construction in the study's species.

3.3 Identification of mRNA Sequences in the Species and Phylogenetic Tree Construction

After domain analysis of ACE2 protein, we found that the most similar species are *Felis catus, Pan troglodytes, Pan paniscus, Equus caballus,* and to characterize this protein and confirm the high similarity of the sequences, we investigated the mRNA sequence of the encoded gene of the protein, and we find that sequence alignment indicates that highly similarity was noticed Fig. 2. Phylogenetic tree of the alignment sequences showed a strong relationship between the species illustrated in the low structure differences in the evolutionary analysis (not more than 0.05) for the defined species Fig. 3.

3.4 Expression Analysis in Human Cells and Tissues

Gene expression is the process by which genetic information is used to create a functioning gene product. These products are often proteins, owing to the importance of this gene in the human body contributes to being entranced for important vulnerable viruses. Because the protein syntheses measure the gene functions, more than 79 samples of different tissues used microarray analysis of tissue mRNA. Data were obtained from the BioGPS database of gene expression patterns in body-specific organs and tissues.

ACE2 shows diversity in conservation according to the protein diversity, normal levels were detected in several body tissues, but the high rate of were detected in the brain, heart, small intestine, colon, kidney, thyroid, breast, testis, and lung.

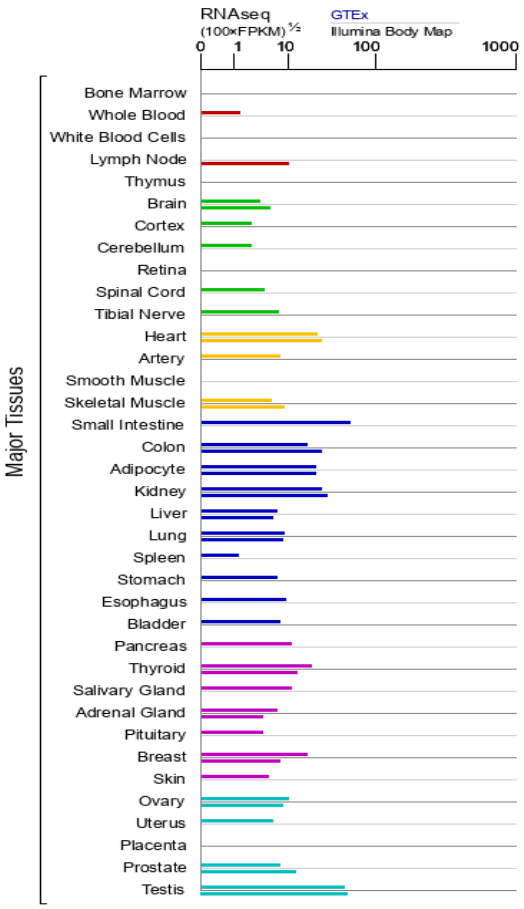


Fig. 4: Microarray data of gene expression levels of ACE2 in the human body organs and tissues. The expression review was collected from the BioGPS database, which included samples of normal human tissues. More than 46 human tissue samples were examined.

3.5 Interaction Network Constructions

Certain genes that interact and collaborate to execute a role modulate the function and activity are represented in the interaction network. ACE2 interaction network is broadly distributed through various families of genes of CALM1, CALM2, CALM3, AAMP, CAT, ISYNA1, AGT, NTS, GHRL, HTR3A, HTR3C, NRP1, HTR3B, DLEU1, and CHRNA10, which have different types of roles ranging from transcriptional activators or co-activator to intracellular signal transducer and transcriptional modulator explained in Fig. 5.

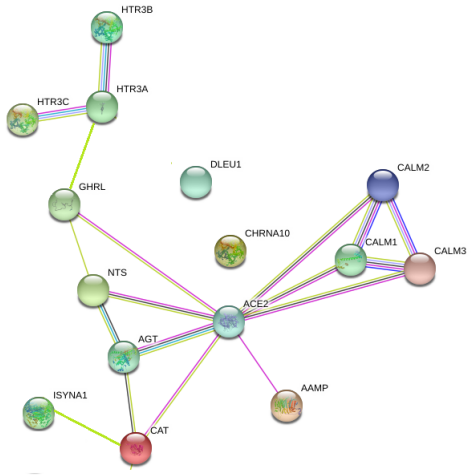


Fig. 5: Gene interaction networks of ACE2. Data were obtained from Gene Card (human gene database) for the definite interacting genes.

4 Discussion

This study investigated seventy-three species of animals that expressed ACE2 receptors on their surface of cells to analyze its structure, function, similarity, differences, relationships between species, hypo, and hyperactivity of the expressed protein based on bioinformatics methods. The data were collected from different reviewed databases based on previously published papers; analysis was done using bioinformatics programs MEGA6 program data sets. Phylogenetic tree, protein domains, expression analysis, and interaction network were explained for the protein sequences. The data showed that phylogenetic analysis of the relation between pan troglodytes, Macaca mulatta, Felis catus, Procyon lotor, and Pan paniscus compared to Homo sapiens shows a high similarity also suggests the high similarity of the function of the gene or the protein. This can be an important result also recommended by Rendon-Marin *et al.,* who suggested that hypothesis based on the analysis of the dimensional protein structure of those animals ACE2 protein [24]. Analysis of the mRNA confirms the previous analysis of the phylogenetic tree of the protein sequence. Also, the mRNA sequences were analyzed with the same method and confirmed the data obtained previously by protein sequences. The results of phylogenetic trees are confirmed by the analysis of the conserved domain for the selected species. We found that *Felis catus, Pan troglodytes, Pan paniscus, and Equus caballus* share with *Homo sapiens* more than 95% or more of the conserved protein domain. This confirms the hypothesis of the similarity between *Felis catus, Pan troglodytes, Pan paniscus, and Equus caballus*. The similarity of the structure reinforces the possibility of function relativity, which is based on genetic structure similarity for proteins and genes. This hypothesis is supported by Wang *et al.,* who suggested that the similarity of the structure may confirm the similarity of the function [25]. These species can be expressed as the protein with the same structure in their cell surface with the same properties as *Homo sapiens*, leading us to the idea of being an actual carrier or transmitter of SARS-COV-2 [26]. Expression analysis of ACE2 shows that specific organs in the body expressed the gene more than other places like the lung, heart, small intestine, kidney, and colon. This may open an important question for the quantity of the virus inside the body, places of its attack, and the possibility of damage to other parts of the body open new venue to study the damage caused by the virus and the ways to avoid it. This is discussed by Mao *et al.,* who support the theory of the causes of the death for unknown reasons while the death suddenly while we do not know the actual reason for the death or which organ is affected especially the complements and symptoms are similar to other diseases like Nepah, Noro or Spanish influenza [27,28]. On the other hand, the infection can transmit when dealing with the dead human body or animal after sacrifices or die. This is consistent with the hypothesis of Park *et al.,* who supports this theory of infection transmission between different species [29,30]. It may reveal with a possibility of infectivity from Homo sapiens to the animal (zoonosis) or animal to *Homo sapiens* (reverse zoonosis), or it may be a possible transmitter to the infection in a definite way which is consistent with the hypothesis of Dhama *et al.* who reviewed the role of SARS, MERS and compare it with the SARS-Cov2 infectivity and transmission ability form animals to humans [31].

The interaction network analysis showed a connection between the responsible gene and other genes in the body works together to complete the function opening another question if the gene of this protein is not the only gene responsible for the increase of the infectivity of the virus and the aggressiveness of the infection or the ability of healing or the random dead also of the ordinary symptoms. Daly *et al.* suggested this, like the neuropilin-1 (NRP1) may act as a host cofactor and enable viral entry may significantly role in the infectivity of the SARS-COV2 [12].

5 Conclusions

We made a systemic analysis to provide more knowledge about this protein's role in the body, providing some possible reason for high and aggressive infection of the virus to open new prospects to further research that can help control the infectivity of this pandemic spread around the world. After this analysis, we can predict that the infectivity of SARS-

COV-2 is not be specialized for Homo sapiens, but there are more than 25 living species, although without any symptoms that can be an infection transmitter to or from Homo sapiens. This is possibly a reason for the widespread SARS-COV-2 worldwide, and the interaction between humans and these species may cause the infection to be increased.

Conflict of interest: The authors have no conflicts of interest to disclose.

References

[1] T.P. Velavan and C.G. Meyer, The COVID-19 epidemic, *Tropical Medicine & International Health.*, **25**, 278–280, 2020.
[2] W.H. Organization. Coronavirus disease (COVID-19). 2020.
[3] W.K. Essa, S.A. Yasin, I.A. Saeed and G.A.M. Ali, Nanofiber-Based Face Masks and Respirators as COVID-19 Protection: A Review, *Membranes.*, **11**, 250-263, 2021.
[4] A.B. Docherty, E.M. Harrison, C.A. Green, H.E. Hardwick, R. Pius, L. Norman, K.A. Holden, J.M. Read, F. Dondelinger and G. Carson, Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study, *BMJ.*, **369**, 1-12, 2020.
[5] M.S.A. Parvez, M.M. Rahman, M.N. Morshed, D. Rahman, S. Anwar and M.J. Hosen, Genetic analysis of SARS-CoV-2 isolates collected from Bangladesh: Insights into the origin, mutational spectrum and possible pathomechanism, *Computational Biology and Chemistry.*, **90**, 107413, 2021.
[6] X. Cao, COVID-19: immunopathology and its implications for therapy, *Nature Reviews Immunology.*, **20**, 269-270, 2020.
[7] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire and D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, *Cell.*, **181**, 281-292, 2020.
[8] Y. Zhang, N. Zheng, P. Hao, Y. Cao and Y. Zhong, A molecular docking model of SARS-CoV S1 protein in complex with its receptor, human ACE2, *Computational Biology and Chemistry.*, **29**, 254-257, 2005.
[9] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.-H. Wu and A. Nitsche, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, *Cell.*, **181**, 271-280, 2020.
[10] C.G. Ziegler, S.J. Allon, S.K. Nyquist, I.M. Mbano, V.N. Miao, C.N. Tzouanas, Y. Cao, A.S. Yousif, J. Bals and B.M. Hauser, SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues, *Cell.*, **181**, 1016-1035. e1019, 2020.
[11] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo and Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, *Science.*, **367**, 1444-1448, 2020.
[12] J.L. Daly, B. Simonetti, K. Klein, K.-E. Chen, M.K. Williamson, C. Antón-Plágaro, D.K. Shoemark, L. Simón-Gracia, M. Bauer and R. Hollandi, Neuropilin-1 is a host factor for SARS-CoV-2 infection, *Science.*, **370**, 861-865, 2020.
[13] K. Tamura, G. Stecher and S. Kumar, MEGA11: Molecular Evolutionary Genetics Analysis version 11, *Molecular Biology and Evolution.*, **38**, 3022–3027, 2021.
[14] B. Chowdhury and G. Garai, A review on multiple sequence alignment from the perspective of genetic algorithm, *Genomics.*, **109**, 419-431, 2017.
[15] J. Euesden, C.M. Lewis and P.F. O'Reilly, PRSice: polygenic risk score software, *Bioinformatics*, **31**, 1466-1468, 2015.
[16] S. Abadi, D. Azouri, T. Pupko and I. Mayrose, Model selection may not be a mandatory step for phylogeny reconstruction, *Nature Communications.*, **10**, 1-11, 2019.
[17] K. Tamura, G. Stecher, D. Peterson, A. Filipski and S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, *Molecular Biology and Evolution.*, **30**, 2725-2729, 2013.
[18] N. Saitou and M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Molecular biology and evolution.*, **4**, 406-425, 1987.
[19] E. Zuckerkandl and L. Pauling. Evolutionary divergence and convergence in proteins. In *Evolving genes and proteins*, Elsevier., 97-166, 1965.
[20] S. Nayfach, A.P. Camargo, F. Schulz, E. Eloe-Fadrosh, S. Roux and N.C. Kyrpides, CheckV assesses the quality and completeness of metagenome-assembled viral genomes, *Nature Biotechnology.*, 1-8, 2020.
[21] A. Garg, N. Singhal, R. Kumar and M. Kumar, mRNALoc: a novel machine-learning based in-silico tool to predict mRNA subcellular localization, *Nucleic Acids Research.*, **48**, W239-W243, 2020.
[22] D.K. Singh, S. Mehra, S. Chatterjee and R.S. Purty, In silico identification and validation of miRNA and their DIR specific targets in Oryza sativa Indica under abiotic stress, *Non-coding RNA research.*, **5**, 167-177, 2020.
[23] M. Bakhtiarizadeh and B. Hosseinpour, In silico Discovery of Conserved and Novel miRNAs from Expressed Sequence Tags in the Chicken (Gallus gallus), *Journal of Agricultural Science and Technology.*, **22**, 667-678, 2020.
[24] S. Rendon-Marin, M. Martinez-Gutierrez, G.R. Whittaker, J.A. Jaimes and J. Ruiz-Saenz, SARS CoV-2 Spike protein in silico interaction with ACE2 receptors from wild and domestic species, *Frontiers in Genetics.*, **12**, 27, 2021.
[25] A. Panzetta, A preliminary study on a new similarity measure for phylogenetic trees, Università Ca'Foscari Venezia., 2016.
[26] P. Bonilauri and G. Rugna, Animal Coronaviruses and SARS-COV-2 in animals, what do we actually know?, *Life.*, **11**, 123, 2021.
27] D.T. Jamison, Disease Control Priorities: improving health and reducing poverty, *The Lancet.*, **391**, e11-e14, 2018.

[28] R. Mao, A. Moa and A. Chughtai, The Epidemiology of Unknown Disease Outbreak Reports Globally, *Global Biosecurity.*, **1**, 2020.
[29] Y.-I. Kim, S.-G. Kim, S.-M. Kim, E.-H. Kim, S.-J. Park, K.-M. Yu, J.-H. Chang, E.J. Kim, S. Lee and M.A.B. Casel, Infection and rapid transmission of SARS-CoV-2 in ferrets, *Cell Host & Microbe.*, **27**, 704-709, 2020.
[30] Y.J. Park, Y.J. Choe, O. Park, S.Y. Park, Y.-M. Kim, J. Kim, S. Kweon, Y. Woo, J. Gwack and S.S. Kim, Contact tracing during coronavirus disease outbreak, South Korea, 2020, *Emerging Infectious Diseases.*, **26**, 2465-2468, 2020.
[31] K. Dhama, S.K. Patel, K. Sharun, M. Pathak, R. Tiwari, M.I. Yatoo, Y.S. Malik, R. Sah, A.A. Rabaan and P.K. Panwar, SARS-CoV-2 jumping the species barrier: zoonotic lessons from SARS, MERS and recent advances to combat this pandemic virus, *Travel Medicine and Infectious Disease.*, **37**, 101830, 2020.